Rats can learn a probability discrimination based on previous trial outcomes in partial reward schedules

PATRICK E. CAMPBELL
Wright State University, Dayton, Ohio 45435

WENDY B. CAMPBELL
Air Force Human Resources Laboratory, Wright Patterson Air Force Base, Ohio 45433

and

BRIAN M. KRUGER and PATRICIA ROBERTS
Wright State University, Dayton, Ohio 45435

In each of two experiments, rats were trained in a runway with either an alternating or an irregular schedule of partial reward. One-third of the animals within each schedule were trained with a small, medium, or large reward. In Experiment 1, the irregular schedule was constructed in such a way that the probability of an R trial following an N trial was .75 and the probability of an N trial following an R trial was also .75. Of course, in an alternating schedule the probabilities are 1.0. The rats learned to run rapidly following N trials and slowly following R trials when trained under either schedule. This effect, however, was greater in the alternating condition and was facilitated by larger rewards. In Experiment 2, the irregular schedule was constructed so that the occurrence of R or N trials was not a reliable predictor of the following trial. Under these conditions, differential running speeds did not develop in the irregular schedule, but the rats in the alternating schedule performed as in Experiment 1. The data are discussed within the context of probability learning and are related to methodological issues of importance to sequential trial experiments.

When extinction performances following alternating or irregular schedules of 50% partial reinforcement are compared following relatively brief training, the alternating schedule results in greater persistence (Capaldi & Hart, 1962), but after extended training the opposite result is found (Capaldi, 1958). This interaction between the type of partial reward schedule and the amount of training has been interpreted as reflecting the importance of two intertrial variables (Capaldi, 1967). With brief training, the number of nonreinforced to reinforced transitions (N-R transitions) should be the dominant factor, while after extended training, the N-length factor should become more important. This theory, of course, accounts nicely for the observed interaction. Irregular schedules have longer N lengths than alternating schedules, but the irregular schedule has fewer N-R transitions.

Campbell, Knouse, and Wroten (1970) suggested that an additional variable was operating to produce the difference in extinction performance between alternating and irregular schedules after extended training. Rats given extended training on alternating schedules typically learn to run rapidly on trials follow-

ing nonreward (TFN) and slowly on trials following reward (TFR). Campbell et al. (1970) hypothesized that the development of such a pattern discrimination should cause alternating schedules to result in less persistence than irregular schedules even when both schedules had a maximum within-day N length of one. Their data supported the hypothesis and led to the suggestion (Campbell, Crumbaugh, Rhodus, & Knouse, 1971) that any variable that facilitated the development of a pattern discrimination should result in a decrease in extinction performance.

The original purpose of the experiments reported here was to test the effects of reward magnitude. In Experiment 1 half of the rats were given extended training on alternating schedules with a small, medium, or large reward. Development of a pattern discrimination was expected to be facilitated by larger rewards, as reported by Campbell et al. (1971), and persistence during extinction was expected to decrease with larger rewards. The other half of the rats were given extended training on irregular schedules with small, medium, or large reward. For these patterning control groups we expected no patterning and expected to confirm the frequent observation that large rewards increase persistence during extinction following training with irregular schedules (cf. Hulse, 1958; Wagner, 1961).

Since no one has reported that rats can learn a pattern discrimination on irregular schedules, we did not expect one to develop. Instead we expected to be relatively uninterested in the acquisition results from the irregular schedules. The fact that we did observe a pattern discrimination with the irregular schedules reduces the relevance of the present extinction data to our original interests. As a result, only the acquisition data are presented.

EXPERIMENT 1

Method

The apparatus consisted of a 216 x 15 x 9 cm runway that was painted flat gray throughout and covered with hinged clear plastic tops. The first 25-cm section served as the starting area, and the final 50-cm section was an L-shaped goalbox. The goalbox contained a 2.5 x 2.5 cm food cup and was separated from the runway by a guillotine door. A masking noise (white) of approximately 80 dB (SPL) was present. Three 25-W red lamps, mounted approximately 1 m over the runway, provided the only available light. Start, run, and goal times to .01 sec were measured over successive 25-, 135-, and 25-cm sections of the alley.

The subjects were 72 experimentally naive male albino 90-day-old rats of the Sprague-Dawley strain. During a 7-day pretraining period, the rats were reduced to 85% of their free-feeding body weights and were familiarized with the runway. Feedings of amounts to be experienced during training were also given in pretrainers similar to the goalbox.

Following pretraining, all animals were given five trials a day for 26 days. One-half of the rats were trained with four irregular partial reward schedules, that is, NRNRN, RNRNR, RNNRR, and NRRNN. These schedules were repeated in sequence every 4 days. The remaining one-half were trained with two alternating partial reward schedules, that is, NRNRN and RNRNR. One-third of the animals within each schedule condition were trained with a small reward (two 45-mg Noyes pellets), a moderate reward (8 pellets), or a large reward (16 pellets).

The animals were tested in squads of six, with one rat from each treatment. The squads were tested in the same order and at the same time each day. The order within each squad was determined randomly each day. All animals were detained in the goalbox for 60 sec on all nonrewarded trials and until the pellets were consumed on all rewarded trials. Each rat ran all five trials successively, and the intertrial interval was held constant at 20 sec. The response times for the first trial of each day were not recorded, and the remaining four trials were recorded as trials following reward (TFR) or as trials following nonreward (TFN). During all trials the rats were given 30 sec to traverse each section of the runway. In the event that this criterion was exceeded, the rats were moved directly to the start of the next segment and a response time of 30 sec was recorded for the previous section.

Results

The response times for each alley section and a total alley measure were converted to speed scores (centimeters per second). These scores were then separated into TFR and TRN trials. Since the individual alley section measures were quite similar to the total alley measure, only the results of the total measure are presented here. Figure 1 shows the total alley speeds for each combination of reward schedule and reward magnitude. An analysis of variance showed a significant (p < .05 for all tests) TFR vs. TFN main effect [F(1,36) = 66.57]. In addition, the TFR vs. TFN discrimination
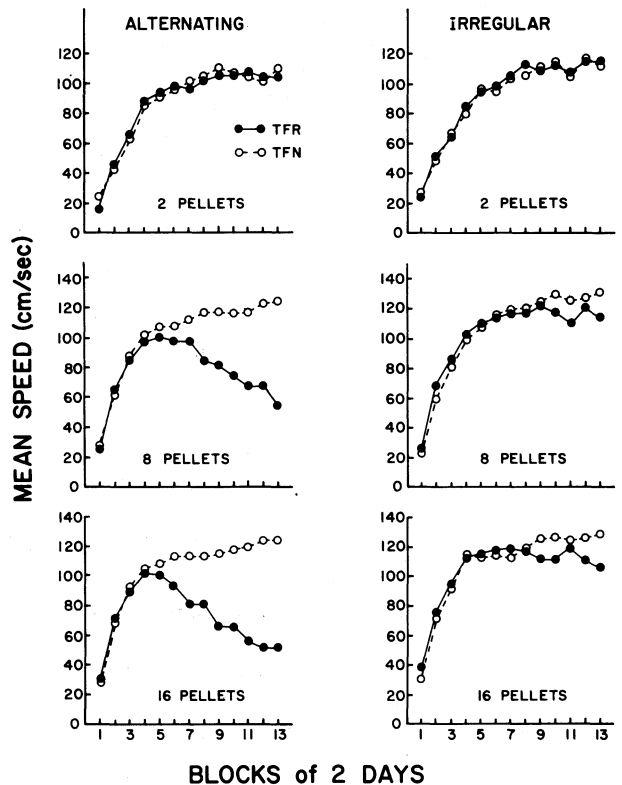

Figure 1. Mean total alley speeds for TFR and TFN trials for all groups in Experiment 1.

developed over trials [F(12,432) = 35.61] was more pronounced in the alternating schedule than in the irregular schedule [F(1,36) = 98.79], and was better the larger the magnitude of reward [F(2,36) = 39.04]. A three-factor interaction combining TFR vs. TFN, reward schedule, and reward magnitude was significant [F(2,36) = 29.13], as was a four-factor interaction combining these factors with trial blocks [F(24, 432) = 3.50]. These results indicate that a pattern discrimination (TFR vs. TFN) developed over trials. The discrimination was more pronounced in the alternating schedule and was increased by larger rewards. This entire pattern of results developed over trials.

The TFR vs. TRN simple-simple main effects were calculated for all groups. First, consider the alternating schedules. Reliable discriminations were found for both 8 pellets [F(1,36) = 156.82] and 16 pellets [F(1,36) = 202.24] but not for 2 pellets [F(1,36) = 1.67]. A similar pattern occurred with the irregular schedules. A discrimination developed for both 8 pellets [F(1,36) = 2.90] and 16 pellets [F(1,36) = 4.17] but not for 2 pellets [F(1,36) = 1.01].

Discussion

The observation that larger rewards facilitate the development of a pattern discrimination in alternating schedules is consistent with the available literature (Campbell et al., 1971). The fact that a previous-trial-outcome discrimination can develop in an irregular schedule is of much greater interest. In

a truly random irregular schedule the occurrence of an N or R trial on Trial n would be of no value in predicting the scheduled outcome of Trial n + 1. The irregular schedules used were designed to minimize the differences between the alternating and irregular schedules on the major sequential trial variables of N length and N-R transitions. Both contained equal numbers of N and R trials, and nearly all of the N-R transitions in both schedules involved an N length of one. As a result of these constraints, our irregular schedules contained probabilistic information about the occurrence of N and R trials, that is, p(N/R) = .75, p(N/N) = .25, p(R/N) = .75, and p(R/R) = .25. Our irregular schedule, then, constituted a probability learning discrimination problem. The finding that rats are sensitive to reinforcement probability or frequency is not new. What is interesting about the present results is that rats can make a probability discrimination when the only apparent source of discriminative stimuli is the outcome of the previous trial.

The observation that pattern discriminations improve with the magnitude of reward is consistent with a number of potential explanations. According to Capaldi (1967), large reward should result in better discrimination because: (1) the aftereffects of large reward are easier to discriminate from nonreward than are the aftereffects of smaller rewards. (2) larger rewards result in stronger habit strength conditioned to the stimulus aftereffects of nonreward; and (3) larger rewards might result in greater inhibition associated with the aftereffects of rewarded trials.

EXPERIMENT 2

We have suggested that the pattern discrimination in Experiment 1 can be considered as a probability discrimination problem, with memories of previous-trial-outcomes serving as the discriminanda. It is possible, of course, that stimuli other than those related to the animal's memory were operative. For example, if differential cues were in any way provided during preparation for rewarded and nonrewarded trials, the difference in speeds on TFR trials and TFN trials may be due to the different pretrial cues on N and R trials, respectively, since TFR trials and TFN trials were probabilistically confounded with N and R trials. Of course, any attempt to assess this possibility by analyzing N and R trials from Experiment 1 would be fruitless due to the confounding with TFR and TFN trials.
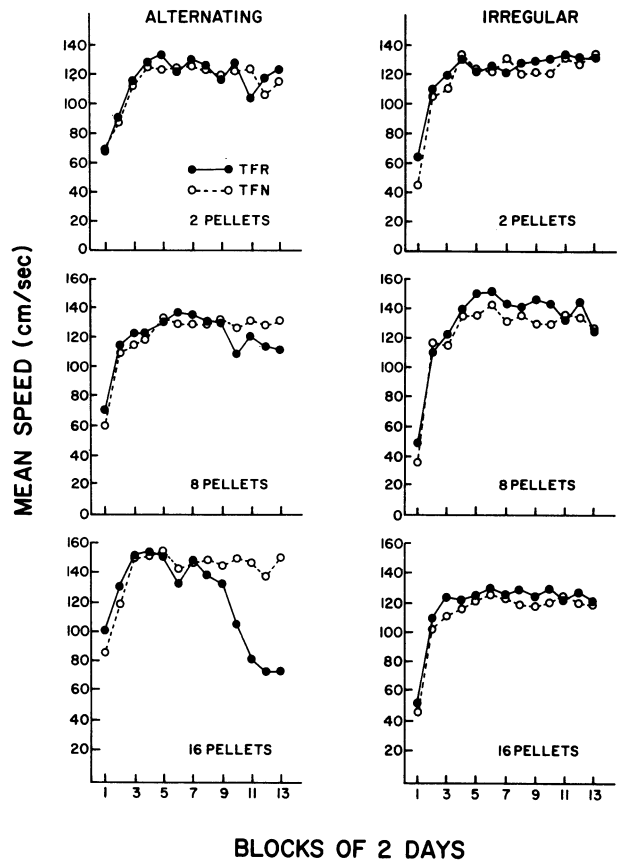
The purpose of Experiment 2 was to repeat Experiment 1 with a new set of irregular schedules. In this experiment, the probabilities of an N or an R trial following an N or an R trial were all .5 for the irregular schedule; therefore, there was no information about the outcome of Trial n based on the outcome of Trial n − 1. Based on this and the absence of any correlation between N and R trials and TFR and TFN trials, it was expected that there would be no difference in speeds on TFR and TFN trials. If any difference was observed, it would be due to the effects of the aftereffects of reward and nonreward per se, rather than being due to differential probabilities of reward on TFR and TFN trials (and speeds on TFR should be faster than on TFN). In order to better assess whether there may have been differential pretrial cues for N and R trials, speeds on N and R trials were also compared.

Method

The subjects were 36 experimentally naive female rats of the Sprague-Dawley strain purchased from Blue-Spruce Farms, Altamont, New York. The animals weighed between 200 and 245 g at the beginning of the experiment.

The 160 x 13 x 9 cm runway was painted flat medium gray throughout and was covered with clear plastic tops. A 4-cm-long goal area of the goalbox contained a food cup and was separated from the 33-cm-long confinement area by a guillotine door. The confinement area was separated from the remainder of the alley by a second guillotine door. Start, run, and goal times were measured to .01 sec for consecutive 25-, 100-, and 25-cm distances, respectively. The first measured distance began 25 cm from the start end of the alley, and the last ended 20 cm inside the confinement area. Two 25-W red lamps located 1 m above the goal area provided the only illumination, and a 75-dB (SPL) white masking noise was present.

The procedural details were the same as in Experiment 1 except as follows. The same two alternating schedules were used (RNRNR and NRNRN), but they were selected at random across days, with the restriction that each schedule be used for 13 days of training. Four new irregular schedules were constructed such that the occurrence of an N or R trial on Trial n did not provide information concerning the occurrence of reward or nonreward on Trial n + 1. The four schedules were selected across days such that each schedule would be used an approximately equal number of times. The four schedules and the number of times each was used were as follows: NNRRN (7 times), RRNNR (7 times), NRRNN (6 times), and RNNRR (6 times).

Figure 2. Mean total alley speeds for TFR and TFN trials for all groups in Experiment 2.

Results

The total alley speed scores were calculated as in Experiment 1 and are shown in Figure 2. An analysis of variance showed a pattern of results quite similar to those of Experiment 1. The results of the simple-simple main effect tests are of greatest interest here. These tests indicated that a significant (p < .05) TFR vs. TFN discrimination developed in two of the alternating schedule groups [F(1,30) = 8.17 for 8 pellets and F(1,30) = 93.21 for 16 pellets] but not in any of the irregular schedule groups [Fs(1,30) ≅ 1.0 for 2, 8, and 16 pellets]. In addition, the speeds on R and N trials for each of the irregular schedules were compared on the final 2 days of training (data not shown). This comparison showed no suggestion of a discrimination between R and N trials.

GENERAL DISCUSSION

It is apparent that rats can learn a probability discrimination based on previous trial outcomes in partial reward schedules. When the various probabilities are certain, as in an alternating schedule, a discrimination is learned relatively easily. Moreover, as Experiment 1 demonstrated, a pattern discrimination can develop even in irregular schedules when the occurrence of an N or R trial provides information about the reward condition scheduled for the following trial. The present data also suggest that pattern discriminations, in either type of schedule, are facilitated by larger rewards.

From a methodological point of view the present data suggest that irregular schedules should be constructed with special attention to the potential informational value of N or R trials. This consideration is likely to be especially important when various constraints are placed on the pattern of trials by sequential trial theory (Capaldi, 1967).

REFERENCES

CAMPBELL, P. E., CRUMBAUGH, C. M., RHODUS, D. M., & KNOUSE, S. B. Magnitude of partial reward and amount of training in the rat: An hypothesis of sequential effects. *Journal of Comparative and Physiological Psychology*, 1971, **75**, 120-128.

CAMPBELL, P. E., KNOUSE, S. B., & WROTEN, J. D. Resistance to extinction in the rat following regular and irregular schedules of partial reward. *Journal of Comparative and Physiological Psychology*, 1970, **72**, 210-215.

CAPALDI, E. J. The effect of different amounts of training on the resistance to extinction of different patterns of partially reinforced responses. *Journal of Comparative and Physiological Psychology*, 1958, **51**, 367-371.

CAPALDI, E. J. A sequential hypothesis of instrumental learning. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 1). New York: Academic Press, 1967.

CAPALDI, E. J., & HART, D. Influence of a small number of partial reinforcement training trials on resistance to extinction. *Journal of Experimental Psychology*, 1962, **64**, 166-171.

HULSE, S. H. Amount and percentage of reinforcement and duration of goal confinement in conditioning and extinction. *Journal of Experimental Psychology*, 1958, **56**, 48-57.

WAGNER, A. R. Effects of amount and percentage of reinforcement and number of acquisition trials on conditioning and extinction. *Journal of Experimental Psychology*, 1961, **62**, 234-242.